Supplementary methods (De la Peña & García-Robles)

HHR motif search methodology

Initial BLAST searches within the *Schistosoma mansoni* genome were done at the Wellcome Trust Sanger Institute (www.sanger.ac.uk) using the whole HHR motifs previously described (Ferbeyre *et al,* 1998). Iterative bioinformatic searches for new HHR motifs among vertebrate genomes were performed through the NCBI-BLAST2 Nucleotide tool at the European Bioinformatics Institute. Seeds for the queries corresponded to U-box/Helix II/P-box motifs (Fig 1A) of naturally-occurring HHR motifs either from previously described metazoans (supplementary Fig S1) or from schistosomal genomes. The number of hits returned by the server usually ranged from just a few hits to several hundreds. Obtained targets were manually inspected to strictly fulfil two initial criteria: (a) Observed changes with respect to the introduced query did not affect to any of the 11 totally conserved nucleotides of the hammerhead ribozyme and (b) changes detected in the targets were either compensatory within the putative helix II or located in loop 2. Filtered hits were analysed for three extra criteria in order to define candidate sequences as HHR motifs (Fig 1B): (c) when the sequence target was extended through 5' and 3' side regions, the presence of Helix I and III had to be found and, (d) Helix I had to be either a loop-closed RNA helix (I/II-type HHR) or an open helix containing an internal loop (II/III-type HHR), and (e) the presence of the expected cleavage site 5'-RUH-3' (where R is a purine residue which has to base pair with the pirimidine residue at the P box, and H can be A, C or U) between Helix I and III. An extra point of validation was the presence and nature of loops 1 and 2, which should establish the tertiary interaction required for *in vivo* self-cleavage (De la Peña *et*

al, 2003; Khvorova *et al,* 2003; Martick and Scott 2006), although due to the known heterogeneity among these interactions (De la Peña *et al,* 2009), a direct confirmation could not be always found. Interestingly, it has to be pointed out that most of the targets already satisfying both the presence of the conserved U and P boxes and allowing the adoption of a typical Helix II usually fulfilled the following criteria (presence of Helix I and III, a RUH cleavage site and loops allowing tertiary interactions), highlighting the high stringency of the search by using just the U-box/Helix II/P-box motifs as initial seeds.Obtained targets fulfilling the requirements were employed as new seeds for BLAST searches, which again, usually resulted in a low number (few hits to few hundreds) of putative HHR motifs that were again manually inspected and selected for further analysis. During the whole search procedure, no selection among the obtained targets was done based on their origin (either RNA/cDNA or genomic DNA source), so the iterative process was followed strictly based on the previous criteria and assuming that all DNA sequence targets may have followed an RNA step.

The occurrence of HH9 was initially detected when used as a seeds sequences from *X. tropicalis* (ie. CTGATGAGCTCCAAGAAGGGCGAAAC corresponding to EST entry DT447185. Fig 2E) or *S. mansoni* (CTGATGAGCTCCAAATAAGACGAAAC corresponding to genomic entry FN357726) among some others. Mammalian seeds derived from HH9 allowed to get HH10 and HHsau.

The quality control of the obtained data from these searches is based in a few major points with its statistically improbable conjunction gives biological significance to the obtained results. Previous bioinformatic searches calculated a probability of chance occurrence of the HHR domain to be about one in 10^{13} nucleotides (Ferbeyre *et al,* 1998), which is two orders of magnitude higher than the current sequence space

available in the last release of the EMBL Nucleotide Sequence Database employed for the searches (Release 102 December 2009; http://www.ebi.ac.uk/embl/Services/DBStats/). Because only naturally occurring U-box/Helix II/P-box motifs were used as queries for the searches, a model bias was obviously introduced in the obtained targets. So, it is the presence in the 5' and 3' surrounding regions of (a) a Helix I, (b) a Helix III and (c) a conserved cleavage site (RUH) not introduced in the search seed, what reassures the significance of the target.

In vitro transcription

Cis-acting hammerheads were synthesized by in vitro transcription of XbaI-linearized plasmids containing the corresponding cDNA inserts immediately preceded and followed by the promotor of T7 RNA polymerase and the XbaI site, respectively. Transcription reactions (20 µl) contained: 40 mM Tris–HCl, pH 8, 6 mM MgCl2, 2 mM spermidine, 0.5 mg/ml RNase-free bovine serum albumin, 0.1% Triton X-100, 10 mM dithiothreitol, 1 mM each of ATP, CTP and GTP, 0.1 mM UTP plus 0.5 µCi/µl [α-32P]UTP, 2 U/µl of human placental ribonuclease inhibitor, 20 ng/µl of plasmid DNA, 4 U/µl of T7 RNA polymerase and 0.1–1 mM of the blocking deoxyoligonucleotide when indicated. Blocking deoxyoligonucleotides were: 33B for HH9 (5'-TTT TTG GAG CTC ATC AGC TGC AGG TAA) and RF-142 for HH5+ (5′-CAT GGA TCT TCA TCA GGA CAC CGA C). After incubation at 37°C for 1 h, products were fractionated by PAGE in 15% gels with 8 M urea and the uncleaved primary transcripts were eluted by crushing the gel pieces and extracting them with phenol saturated with buffer (Tris–HCl 10 mM, pH 7.5, EDTA 1 mM, SDS 0.1%), recovered by ethanol precipitation and resuspended in deionized and sterile water.

Cis cleavage kinetics under protein-free conditions

For determining the cleaving rate constants, uncleaved primary transcripts (from 1 nM to 1 μM) were incubated in 20 μl of 50 mM PIPES-NaOH, pH 6.5 for 1 min at 95°C and slowly cooled to 25°C for 15 min. After taking a zero-time aliquot, self-cleavage reactions were triggered by adding $MgCl_2$ to 1 mM. Aliquots were removed at appropriate time intervals and quenched with a 5-fold excess of stop solution at 0°C. Substrates and cleavage products were separated by PAGE in 15% denaturing gels. The product fraction at different times F_t was determined by quantitative scanning of the corresponding gel bands and fitted to the equation $F_t = F_o+F_\infty(1-e^{-kt})$, where F_o and F_∞ are the product fractions at zero time and at the reaction endpoint, respectively, and k is the first order rate constant of cleavage (k_{obs}).

Supplementary Table S1. Examples of HHR hits obtained in genomic data (ESTs and Genomic sequences) from amphibians and lampreys through iterative bioinformatic searches in the databases (see Materials and methods)

ESTs	***Accession n.***	***Strand***	***Origin***	***Position***
Xenopus Tropicalis	DT447185	Plus	RNA	663
	DT447184	Minus	RNA	594
	DN046758	Minus	RNA	402
	CX851145	Plus	RNA	168
	CX838586	Plus	RNA	150
	CX809771	Plus	RNA	38
	CX804623	Plus	RNA	190
	DN049774	Plus	RNA	100
Ambystoma tigrinum	CN047845	Plus	RNA	87
Ambystoma mexicanum	CO786334	Minus	RNA	155
Cynops pyrrhogaster	FS293747	Minus	RNA	541
Petromyzon marinus	EG023975	Minus	RNA	426
	EB721675	Plus	RNA	310
	FD713511	Plus	RNA	217
	CO547793	Minus	RNA	239
	EG335887	Plus	RNA	75
	EB718691	Minus	RNA	67
	FD720012	Minus	RNA	174
	EC425355	Minus	RNA	100
	EG338286	Plus	RNA	214
	EC426782	Minus	RNA	245
	FD722165	Minus	RNA	729
Lethenteron japonicum	DC612391	Plus	RNA	205
	DC617201	Minus	RNA	226
	DC616422	Plus	RNA	408

Genomic	***Gene***	***Strand***	***Origin***	***Position***
Xenopus tropicalis	RT-like	Minus	DNA	scaffold_375 1057652
	RT-like	Minus	DNA	scaffold_179: 1837564
	Transposase	Plus	DNA	scaffold_68: 1984600
	Telomerase RT (TERT)	Minus	DNA	scaffold_57: 1529678
	Zinc finger protein	Plus	DNA	scaffold_578: 494798

Petromyzon marinus	S.japonicum RT-like	Minus	DNA	Contig6481: 17053
	ElastaseB	Plus	DNA	Contig1157: 9815
	Protease inhibitor	Plus	DNA	Contig19006: 8075
	RecQ DNA Helicase	Plus	DNA	Contig42040: 441
Ambystoma mexicanum	CO786334	Minus	RNA	155
	EU686405	Plus	DNA	34032
	EU686414	Minus	DNA	14206
	EU686401	Plus	DNA	13346
	EU686400	Plus	DNA	50341
	EU686411	Plus	DNA	70802
	EU686412	Plus	DNA	152069 176397 (double)

SI figure legends

Supplementary Figure S1. HHR motifs in metazoans described so far in the literature. Helixes II are shown in red.

Supplementary Figure S2. *In vitro* self-cleavage capabilities of HH9 motif from *Homo sapiens* versus the viroidal HH5+ motif from CChMVd(+). **(A)** RNAs obtained after 10 min transcription of HH9 and HH5+ constructs revealed similar levels of self-cleavage. **(B)** Kinetic analysis of HH9 at 25^{o}C under low pH (6.5) and Mg^{2+} concentration (1 mM) revealed a fast cleavage reaction (k_{obs}=2.43±0.22 min^{-1}), within the same order as the observed one for HH5+ under identical reaction conditions (4.68 ±0.42 min^{-1}) and other naturally occurring HHRs (De la Peña et al, 2003; Khvorova et al, 2003). **(C)** Structural comparison of the loop-loop interactions described for the *S. mansoni* HHR (right. Martick and Scott, Cell 2006) and the deduced ones for the human HH9 (left). Nucleotidic differences between both motifs are shown in grey.

Supplementary Figure S3. Alignment of the intronic region containing the HH9 motif of the RECK genes from 32 endothermic vertebrates. Scheme of the *Homo sapiens* RECK pre-mRNA transcript is shown at the top.

Supplementary Figure S4. Alignment of the intronic region containing the HH10 motif from 19 mammalian C10orf118 genes. Scheme of the *Homo sapiens* C10orf118 pre-mRNA transcript is shown at the top.

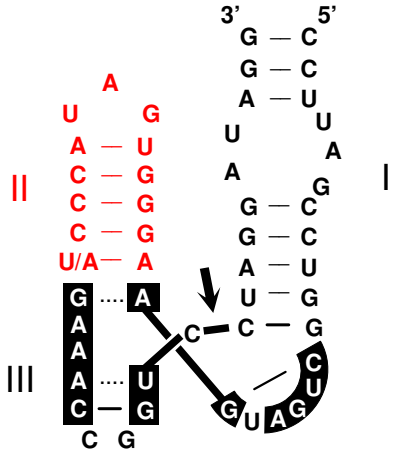
Supplementary Figure S5. ESTs from *Homo sapiens* starting at the 3' end of the conserved region of helix I from HH10 (arrows in red). Changes and deleted positions detected in the ESTs sequences (boxed) with respect human HH10 are shown in black or within brackets. Nucleotides in green correspond to the pCMV6 vector used in the EST cloning procedure.

Supplementary Figure S6. HH10 sequence variability among 21 different mammals. Sequence heterogeneity with respect to the human HH10 version is highlighted in the same color as the name of the species. *Ins* and *Del* refer to inserted and deleted nucleotides respectively. The HH10 case of the American pika *Ochotona princeps*, a lagomorph closely related to rodents, is highlighted within a pink box. Inactivating mutations at the catalytic centre and the Helix I of the ribozyme are highlighted with an asterisk.

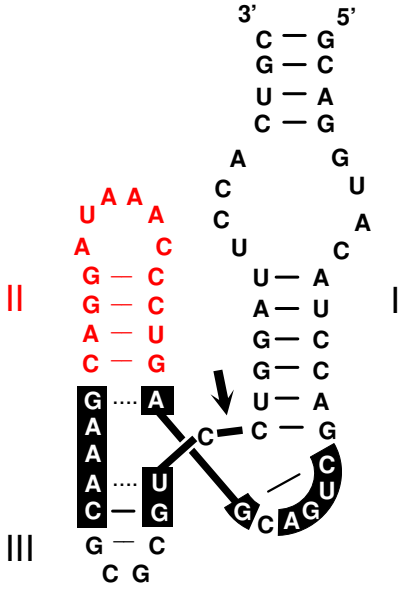
Supplementary Figure S7. HHR motifs found within (left) or at the 5' side (right) of the C10orf118 homologous gene in *S. mansoni*. Sequence differences respect to human HH10 are highlighted in brown color.

Supplementary Fig S1
(De la Peña & García-Robles)

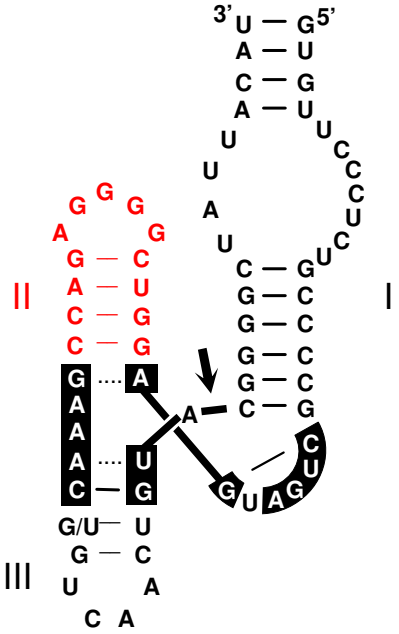
Newts and salamanders



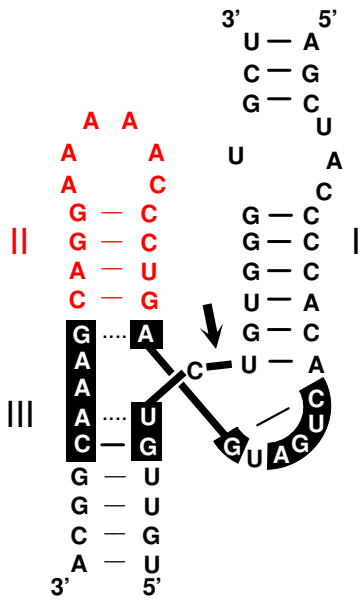
Schistosoma spp.



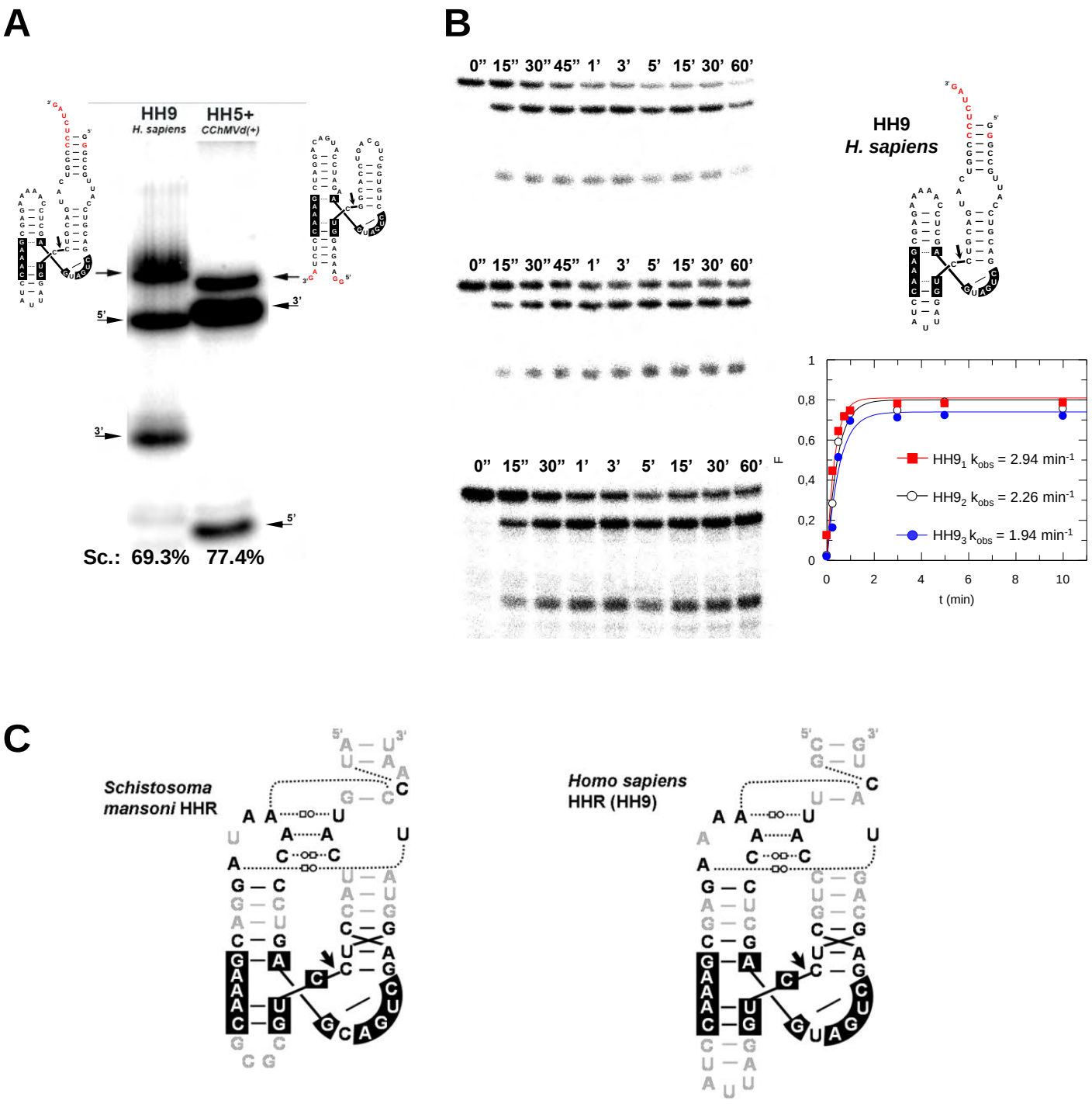
Dolichopoda spp.



Rodents (Clec2)

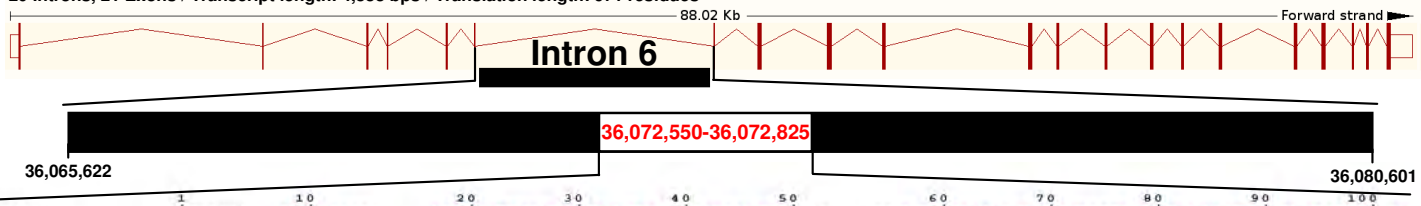


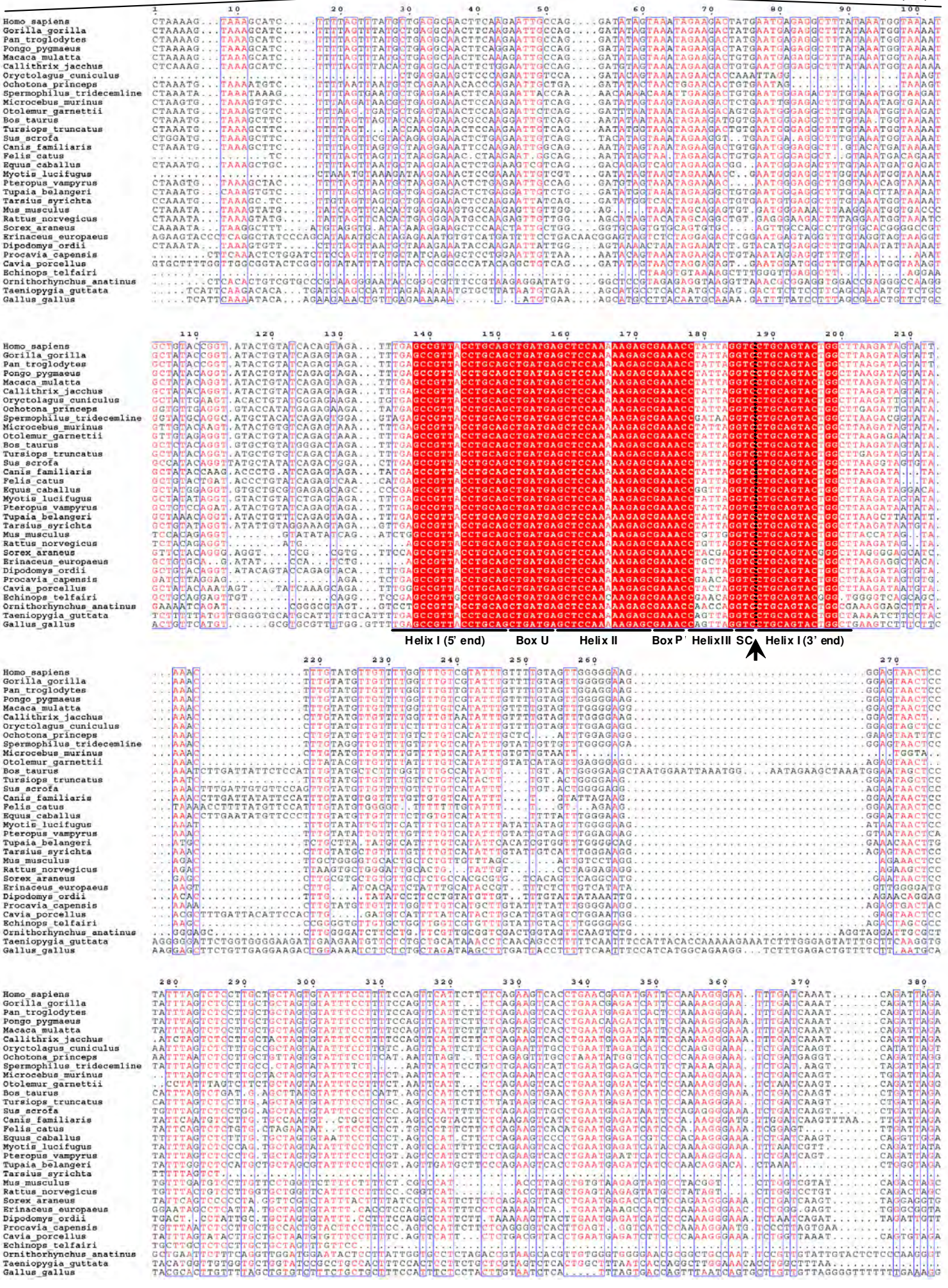
Supplementary Fig S2
(De la Peña & García-Robles)



RECK (Transcript ID: ENST00000377966); *H. sapiens* Chromosome 9, contig NT_008413.18
20 Introns, 21 Exons / Transcript length: 4,888 bps / Translation length: 971 residues

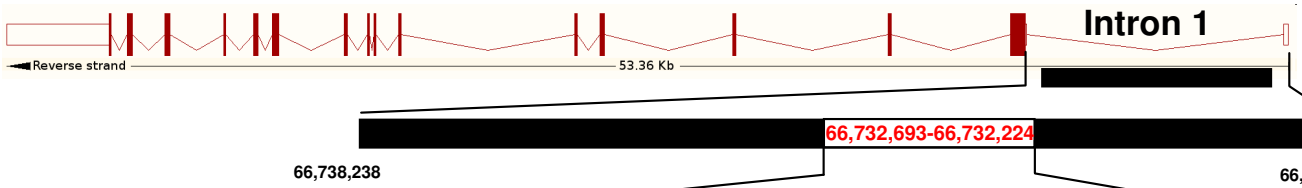
Supplementary Fig S3

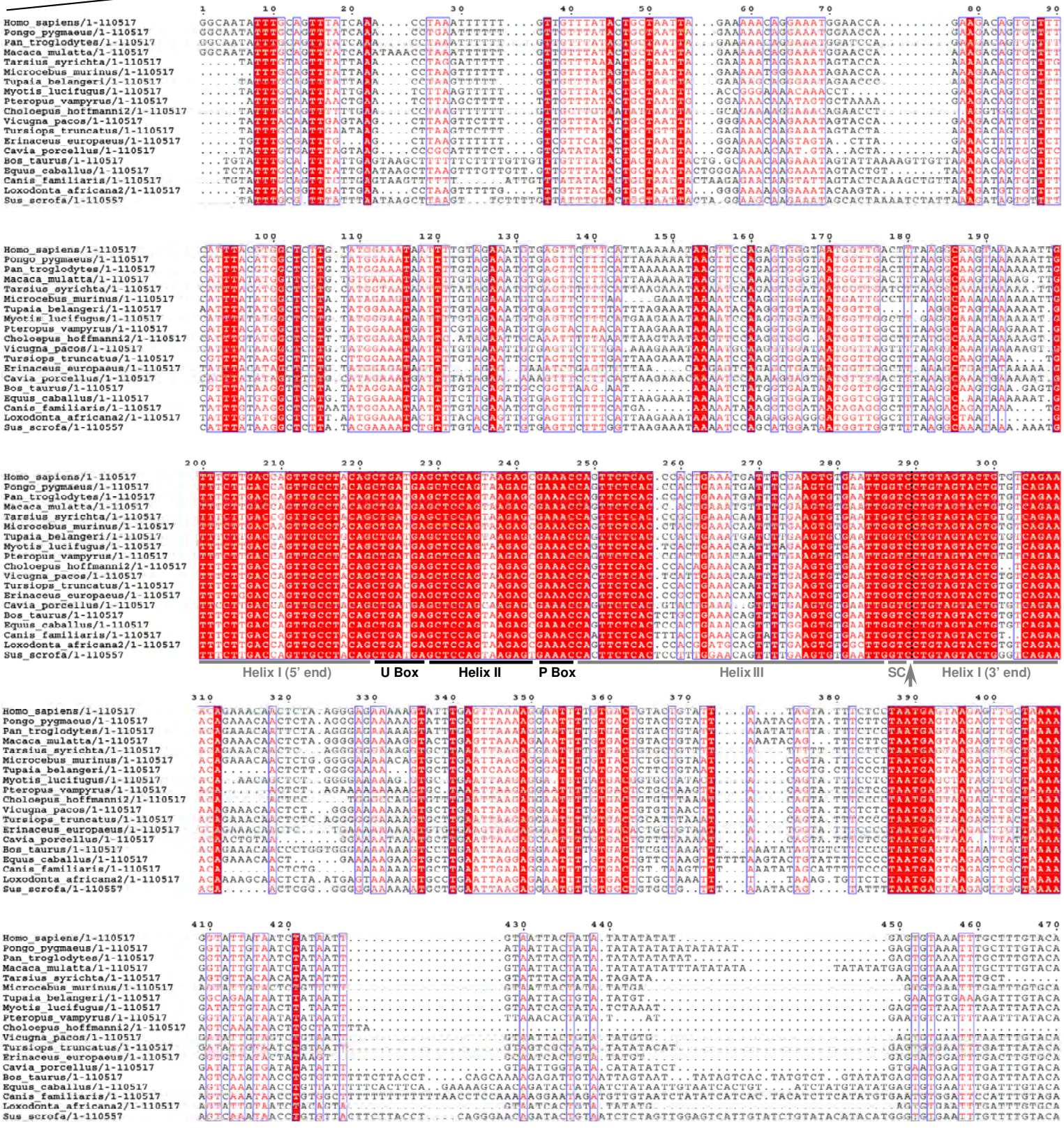




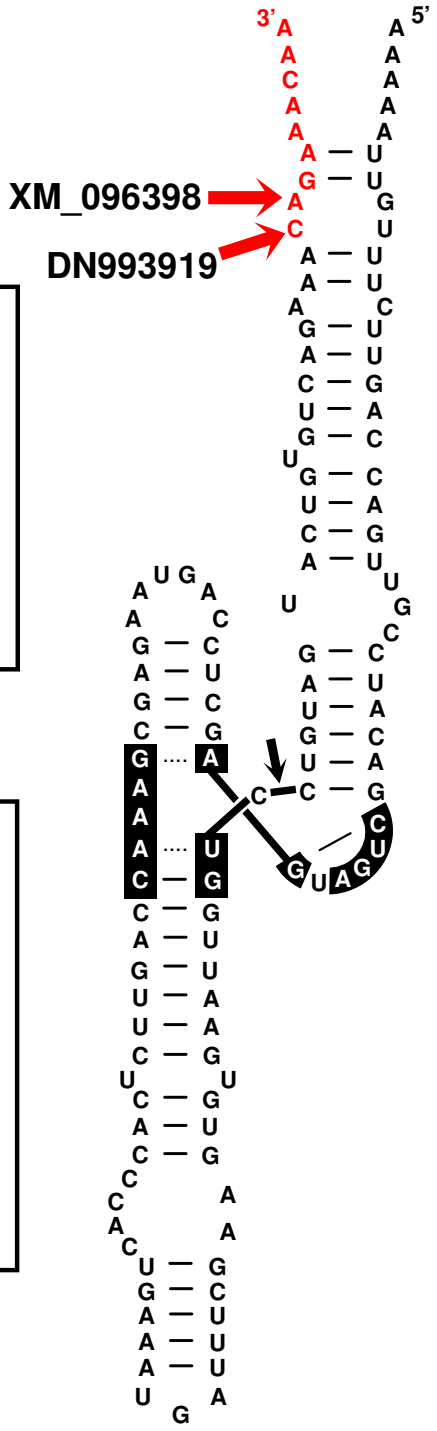
Supplementary Fig S4
(De la Peña & García-Robles)

C10orf118 or CTCL L14-2 (Transcript ID:ENST00000369287); *H. sapiens* Chromosome 10, contig NT_030059.13
15 Introns, 16 Exons / Transcript length: 7,245 bps / Translation length: 898 residues





Supplementary Fig S5
(De la Peña & García-Robles)



>DN993919

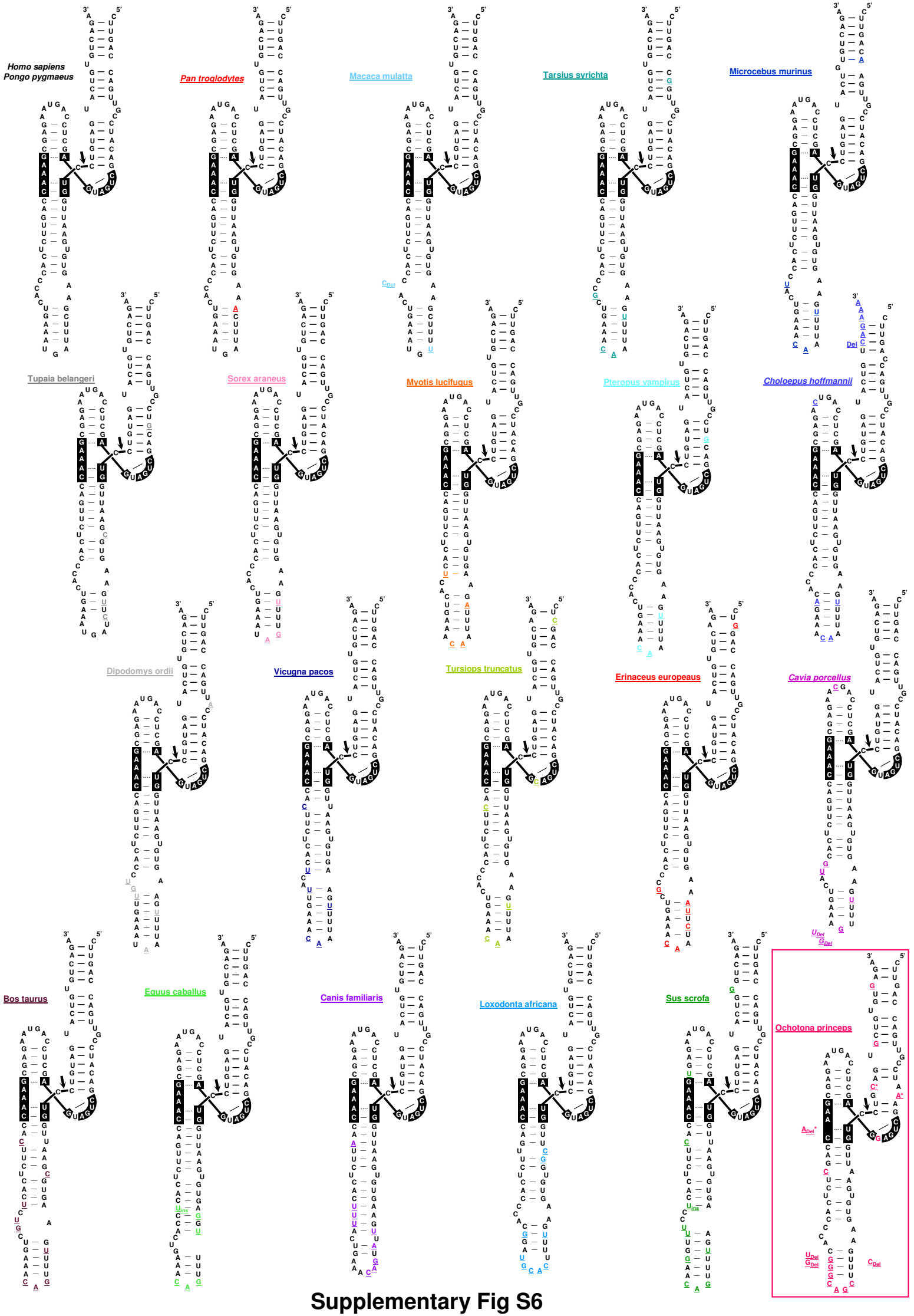
5′CACGAGAA(A)C(A)ACTCTAAGG(G)AGAAAAA(G)TATTTGAGTTAAAAGGAATTTTGTGACTGTACTGTATTATAG…(440 nt. Intron 1)-3′

mol_type: mRNA from Human adult brain
tissue_type: Whole brain

>XM_096389

5′NAAGGGGTCCACAATTGTAAACGACTCACTATAGGCGGCCGCGAAATTCGCACGAGGAACAACTCTAAGGAGAAAAATATTTGAGTTAAAAGGAATTTTGTGACTGTACTGTA…(795 nt. Intron 1)-3′

mol_type: mRNA from *H. sapiens*
tissue_type: Brain
(pCMV6-XL5 vector)

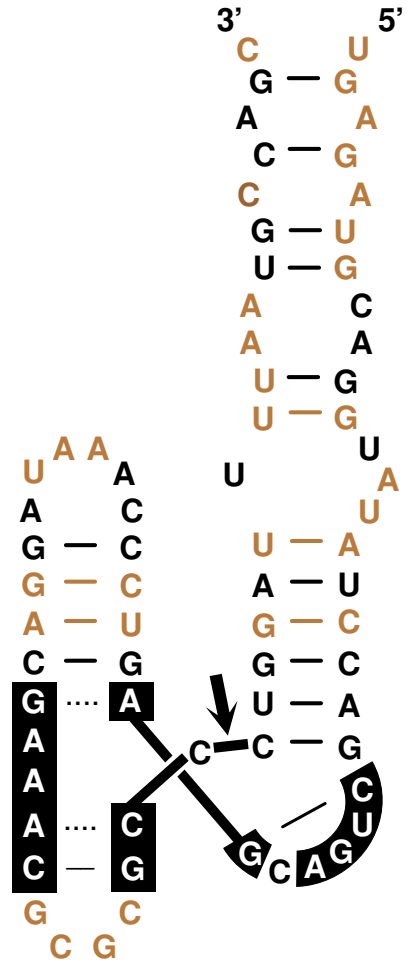


Supplementary Fig S6

Supplementary Fig S7
(De la Peña & García-Robles)

HH10-like (I)

S. mansoni



HH10-like (II)

S. mansoni